DOCUMENT RESUME

ED 065 314 24 SE 014 286

AUTHOR Nelson, Barbara A.; Frayer, Dorothy A.
TITLE The Effects on Short- and Long-Term Retention of Discovery and Expository Methods of Presenting Selected Geometry Concepts: A Replication.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO TR-208
BUREAU NO BR-5-0216
PUB DATE Feb 72
CONTRACT OEC-5-10-154
NOTE 43p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Discovery Learning; Expository Writing; *Geometric Concepts; *Instruction; *Learning; *Research; Retention; *Secondary School Mathematics; Teaching Techniques

ABSTRACT

This study was designed to investigate the effects of discovery and expository methods of presentation on the immediate acquisition and retention of geometry concepts by seventh graders. Four geometry concepts, "quadrilateral," "rhombus," "trapezoid," and "parallelogram," were presented in written lessons which used either an expository or a discovery approach. Ss studied the lessons on four consecutive days. They were given a test on the concepts either immediately after completion of the lessons or one day or 21 days after completion of the lessons. Ss who received the test one day after completion were retested 21 days after completion. The findings of the study were as follows: (1) Ss who studied the expository lessons had significantly higher scores than Ss who studied the discovery lessons; (2) An independent groups analysis revealed no significant difference between treatment groups. However, a repeated measures analysis revealed that Ss in the discovery group tended to forget less than Ss in the expository group over the 21-day retention interval; and (3) The expository group spent only one-third as much time on the lessons, yet had immediate acquisition superior to the discovery group and 21-day retention equal to the discovery group. (Author/JM)

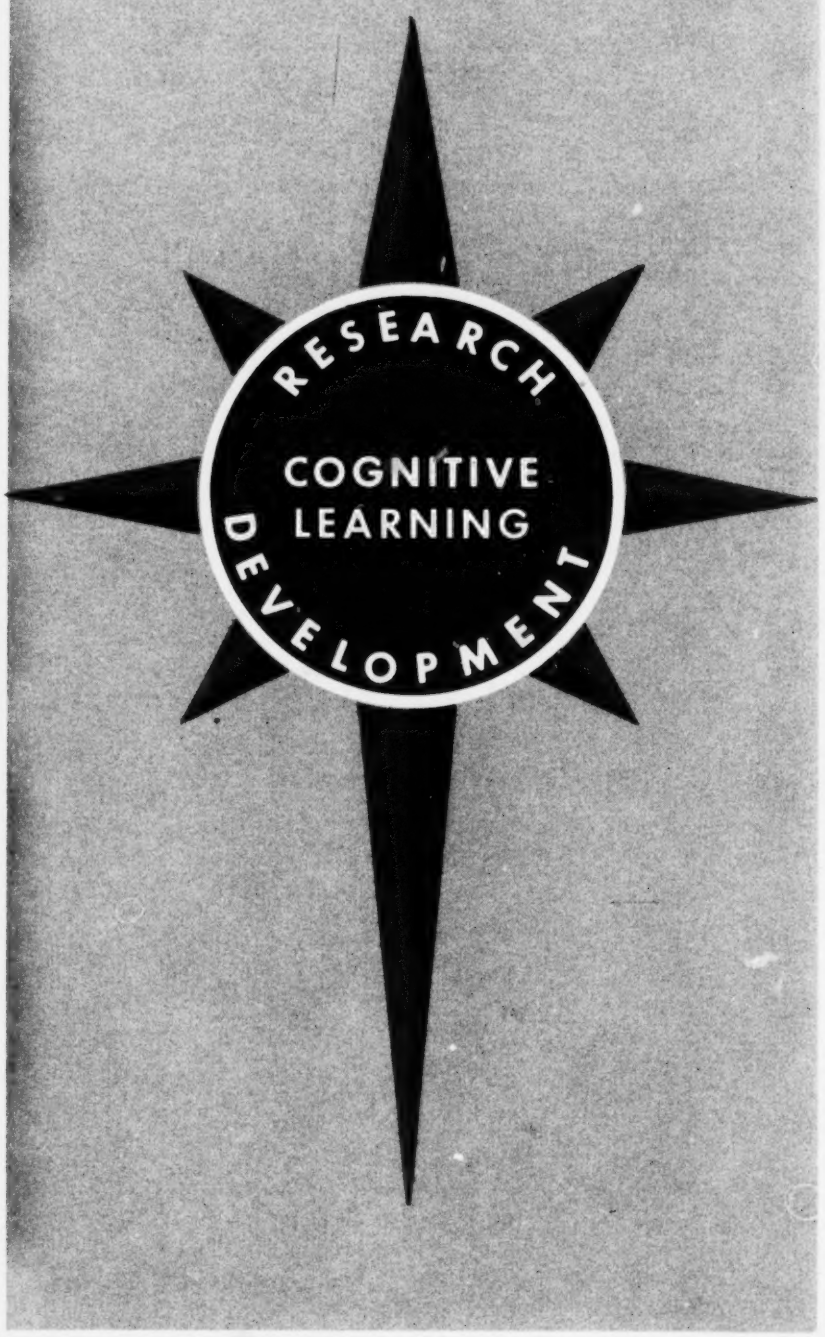
THE EFFECTS ON SHORT- AND LONG-TERM RETENTION OF DISCOVERY AND EXPOSITORY METHODS OF PRESENTING SELECTED GEOMETRY CONCEPTS: A REPLICATION

WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR
COGNITIVE LEARNING

NCERD
BR 5021C
PA 24
SE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY



Technical Report No. 208

THE EFFECTS ON SHORT- AND LONG-TERM RETENTION OF DISCOVERY AND EXPOSITORY METHODS OF PRESENTING SELECTED GEOMETRY CONCEPTS: A REPLICATION

By Barbara A. Nelson and Dorothy A. Frayer

Report from the Conditions of Learning
and Instruction Component of Program 1

Herbert J. Klausmeier, Principal Investigator
Dorothy A. Frayer, Assistant Scientist

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

February 1972

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Project on Variables and Processes in Cognitive Learning in Program 1, Conditions and Processes of Learning. General objectives of the Program are to generate knowledge and develop general taxonomies, models, or theories of cognitive learning, and to utilize the knowledge in the development of curriculum materials and procedures. Contributing to these Program objectives, this project has these objectives: to ascertain the important variables in cognitive learning and to apply relevant knowledge to the development of instructional materials and to the programming of instruction for individual students; to clarify the basic processes and abilities involved in concept learning; and to develop a system of individually guided motivation for use in the elementary school.

Acknowledgments

Many people contributed their time and talents to the completion of this study. Foremost among these is Thomas J. Fischbach, whose creativity and statistical expertise greatly strengthened the design of the experiment. Dr. Herbert J. Klausmeier originally conceived the study and offered many perceptive suggestions which helped bring it to fruition. Barbara J. Marten generously provided assistance in revising lessons and developing the lesson and exercises on use of the ruler.

We deeply appreciate the cooperation of the Oregon (Wisconsin) Junior High School and Beloit (Wisconsin) Public Schools in this study. Particularly helpful were Mr. Karl Ahrens, Principal of the Oregon Junior High School, and Joseph W. Rhodes, Assistant Superintendent of the Beloit Public Schools. The principals and teachers of the junior high schools in Beloit contributed greatly to the second pilot study and the main study by meeting our many needs for assistance. We especially appreciated their interest in improving education through research. A special thanks goes to Mrs. Joyce Fluck and Mrs. Luella Pohlman who conscientiously and capably administered the lessons and tests for the main study.

A study of this magnitude cannot be completed without the behind-the-scenes help of many people who often go without recognition. We sincerely appreciate the assistance of the duplicating shop workers, typists, keypunchers, and computer programmers at the Center whose standards of excellence are prime factors in good research.

Contents

		Page
	Acknowledgments	iv
	List of Tables	vii
	List of Figures	vii
	Abstract	ix
I.	Introduction	1
II.	Method	5
	Pilot Study I	5
	Subjects	5
	Materials	5
	Procedure	6
	Results	6
	Pilot Study II	7
	Subjects	7
	Materials	7
	Procedure	7
	Design	7
	Results	8
	Main Study	8
	Subjects	8
	Materials	8
	Procedure	8
	Design	10
III.	Results	11
	Pilot Study II	11
	Main Study	13
	Psychometric Characteristics of Test BN	13
	Analysis of the Data	13
IV.	Discussion and Conclusions	17
	Discussion	17
	Conclusions	18
	References	21
	Appendices	
	A. Teacher Questionnaire: Knowledge of Geometry Concepts	23
	B. Instructions to Students	27
	C. Arithmetic Achievement Scores and Scores on Test BN for Each Subject	31

List of Tables

Table		Page
1	Schedule for Pilot Study I	7
2	Schedule for Pilot Study II	8
3	Schedule for the Main Study	9
4	Number of Subjects in Each Treatment by Achievement Group	10
5	Number of Subjects and Observed Means and Standard Deviations of Multiple Choice, Production, and Total Scores for Treatment Groups by Class in Pilot Study II	11
6	Analysis of Variance of Subtest Scores on Forms BN-1 and BN-2 in Pilot Study II	12
7	Number of Subjects and Observed Means and Standard Deviations of Scores for Treatment Groups by Achievement Level in the Main Study	13
8	Analysis of Variance of Total Scores on Test BN in the Main Study	14
9	Mean Scores and Changes in Scores for Repeated Administrations of the Same Test to Discovery, Expository, and Control Groups at 1 Day and 21 Days	15
10	Analysis of Variance of Change Scores in the Repeated Measures Design	16

List of Figures

Figure		Page
1	Mean scores of discovery, expository, and control groups immediately, 1 day, and 21 days after completion of lessons (independent groups analysis)	14
2	Mean scores of discovery, expository, and control groups 1 day and 21 days after completion of lessons (repeated measures analysis)	15

Abstract

This study was designed to investigate the effects of discovery and expository methods of presentation on the immediate acquisition and retention of geometry concepts by seventh graders.

Four geometry concepts, *quadrilateral*, *rhombus*, *trapezoid*, and *parallelogram*, were presented in written lessons which used either an expository or a discovery approach. The expository approach was characterized by the presentation of the name and definition of the concept, followed by positive and negative examples in which the relevant attributes were explicitly pointed out. Under the discovery method, the name of the concept was presented, followed by a series of positive and negative examples. *Ss* were asked to describe the examples and to state how they were alike and how they were different. After all examples were presented, the definition of the concept was given.

Ss studied the lessons on four consecutive days. They were given a test on the concepts either immediately after completion of the lessons or 1 day or 21 days after completion of the lessons. *Ss* who received the test 1 day after completion were retested 21 days after completion.

The findings of the study were as follows:

1. Method of presentation differentially affected immediate acquisition. *Ss* who studied the expository lessons had significantly higher scores than *Ss* who studied the discovery lessons.
2. The effect of method of presentation on retention was not as clear. An independent groups analysis revealed no significant difference between treatment groups. However, a repeated measures analysis revealed that *Ss* in the discovery group tended to forget less than *Ss* in the expository group over the 21-day retention interval.
3. The expository method was more efficient than the discovery method. The expository group spent only one-third as much time on the lessons, yet had immediate acquisition superior to the discovery group and 21-day retention equal to the discovery group.

I
Introduction

Proponents of the discovery method of learning laud its superiority over the more conventional expository method of learning. Wittrock (1966) has summarized some of these claims:

> . . . . learning by discovery produces knowledge which transfers to new situations. Through practice at problem solving it develops problem solving ability. It is intrinsically motivating and is its own reward. By being taught to solve problems, to behave in a scientific and inductive fashion, and to go beyond the data, a student is helped to become a mature person. It is a useful conceptualization for the teaching of many subjects in schools. . . . It is an important end in its own right. . . . One must learn to produce rather than to reproduce answers and knowledge. (P. 36)

These claims involve a combination of vaguely defined dependent variables, cultural values, and intuition. Few of these claims are based on empirical research, however, since little research concerning the relative merits of discovery and expository learning has been generated by the debate. As Morrisett (1966) concluded at a conference on Learning by Discovery, "research on the topic of discovery . . . is relatively impoverished" (p. 179).

Not only have there been few studies on discovery learning, but the results of these studies tend to be contradictory. For example, six studies have compared discovery learning with either rote or expository learning of subject matter taught over periods ranging from 17 days to 7 months. On tests given at the close of the instructional period, McConnell (1934) and Worthen (1968) found expository learning superior, Thiele (1938) and Swenson (1949) found discovery learning superior, while Anderson (1949) and Nichols (1956) found no difference between discovery and expository learning.

Two outcomes of discovery learning have been more consistently observed. First, discovery and expository learning lead to equal performance on tests given immediately after lessons taught over a period of only one or two days (Ray, 1961; Scott, 1970). Second, discovery learning is superior to expository learning on long-term retention (Ray, 1961; Scott, 1970; Swenson, 1949; Worthen, 1968). These results suggest that retention rather than original learning might be a key advantage of the discovery method.

Unfortunately, methodological problems in many of the studies showing better long-term retention for discovery learning lessen our confidence in the finding. In addition, the results on tests given immediately after long instructional periods, which reflect retention as well as acquisition, have been contradictory. Thus, the effects of discovery learning on retention are not clear-cut. Educators, however, are keenly interested in determining the comparative advantages of expository and discovery learning, since this would have definite implications for instruction. For this reason, the present study was undertaken in an attempt to replicate one of the studies (Scott, 1970) which found better long-term retention for discovery learners.

To provide an understanding of what is meant by "discovery learning" and the methodological problems encountered in studying its effects, a few of the studies which have compared discovery and expository learning will be reviewed. In each case, attention will be focused on how the researcher operationalized the discovery method and what the effects of that method were on immediate acquisition and retention.

One of the earliest studies of discovery learning compared the effects of discovery and rote methods of learning on acquisition. In this study McConnell (1934) presented 100 addition and 100 subtraction facts to second-grade students over a seven-month period. Students who had learned by an authoritative method, in which they were told to memorize the facts, did better on a speeded task than students who had learned by a discovery method in which they were told to discover the generalization involved. The task required giving the answers to addition and subtraction problems presented on flash cards at the rate of one per 4 2/7 seconds.

Thiele (1938) using the same general task and age group obtained different results. Second-grade teachers taught 100 addition and 100 subtraction facts by two methods: generalization, in which students were told to look for a generalization, and drill, which was similar to McConnell's authoritative method. To insure uniformity within methods, instructions for the teachers of each set of experimental classes were carefully prepared and explained at teachers' meetings and each teacher was visited twice during the 15 weeks of instruction. A test consisting of the presentation of 100 one-digit facts on flash cards at the rate of one per 4 seconds was administered twice, as a pretest and posttest. Thiele found that students who learned by the generalization method performed better than students who learned by the drill method. He suggested that this result could be attributable in part to the enthusiasm of the teachers "for a method of teaching the addition facts which removed much of the tedium of repetitive drill..." (p. 77).

A third study using a similar methodology was carried out by Swenson (1949). Classroom teachers taught 100 addition facts to fourth-grade students using one of three methods of instruction. Under the generalization method, teachers presented the addition facts in groups determined by some unifying idea or generalization and encouraged students to look for interrelationships among facts. Under the drill method, teachers presented the addition facts to the students in a random order and discouraged any type of generalization. Finally, under the drill-plus method, an attempt was made to simulate common practices. Students in this group were presented the facts in groups which were organized by the size of the sums and were allowed to verify results by counting and manipulating concrete objects. Each teacher taught only one class, using the experimental method which most closely approximated the method he normally used, as determined by an interview and attitude inventory. Instruction continued for a period of five weeks. Teachers were allowed to adjust time and materials to differences among classes and pupils. On each day the drill teachers spent an average of 11 minutes on instruction and 24 minutes on drill, while the generalization teachers spent 27 minutes on instruction and 18 minutes on drill. A test consisting of 100 addition facts presented at the rate of one every 4 seconds was used to measure immediate acquisition and retention 5 1/2, 8, and 12 weeks later. Swenson found that on immediate acquisition the generalization group was superior to the other two groups, while the drill group was superior to the drill-plus group. These differences were maintained through the three retention intervals. There were, however, no significant differences among methods on retention independent of the initial difference in learning.

Anderson (1949) compared two methods of teaching part of the regular math curriculum to fourth-grade students over a seven-month period. The two methods of teaching were drill and meaning which were comparable to the drill and generalization methods used by Swenson. Each teacher in the study used only one of the methods and was given much freedom in deciding the content, amount of time spent on instruction, and number of repetitions of the number facts. Each was, however, provided with a manual and training in his method. Standardized arithmetic tests given at the end of the year revealed no differences due to teaching method.

In Anderson's study each teacher used only one of the methods being compared. Worthen (1968) attempted to reduce the teacher effect by requiring that each teacher use both methods. Fifth- and sixth-grade students were given six weeks of instruction in elementary mathematics by one of two methods. In the discovery (D) method of presentation the name of each concept was delayed until the end of the instructional sequence; in the expository (E) method the name of the concept was given initially, followed by examples of the concept. Each teacher taught a class by each of the two methods for the six-week period. Teachers were given a summary of model teaching behavior for both methods. Observers' and pupils' ratings confirmed that the teachers were able to vary their behaviors sufficiently to render the two methods different. The same concepts were taught under each of the methods. Materials were equated for the number and type

of examples and for the amount of verbal presentation. The same amount of time was allowed for each method. Worthen found that while the expository group was superior on initial acquisition, the discovery group was superior on retention five weeks later.

Ray (1961) taught high school students the use of the micrometer by one of two methods of instruction. Method A involved direct and detailed instruction in which the teacher presented the material, reviewed important points, and solved examples. The teacher presented the learning material without a break for 40 minutes. Method B involved directed discovery. The students studied material on their own and thought about leading questions asked by the teacher. Out of the 40 minutes, 19 minutes were spent in silence during which the students could discover principles and make generalizations. Tape recordings of all oral instructions assured constancy of conditions among groups. Subjects were ninth-grade students stratified by intelligence level. Ray found that (a) initial learning was the same under both treatment methods, (b) retention at one week was the same, but retention at six weeks favored the directed discovery group, and (c) there was no interaction of teaching method and intellectual ability.

Nichols (1956) compared two methods of presenting plane geometry to college freshmen. *Ss* in the dependence group depended on the teacher to present statements of assumptions, theorems, and definitions, and verbalizations of principles through deductive arguments. The teacher was told what to teach every day but detailed explanations of how to teach the material were not given. Each session was recorded on tape. *Ss* in the structured search group discovered relationships through a series of concrete experiences such as measuring and cutting out drawings of geometric figures. The experiences leading to the discovery of each relationship were structured through written materials which included incomplete verbal statements of principles which *Ss* were to complete. Unlike the *Ss* in the dependence group, the *Ss* in the structured search group worked independently. The two treatment groups were taught on an alternating basis by three teachers to randomize teacher effect. Each teacher taught each method five or seven days during the treatment period. The teacher who was not teaching on a certain day served as an observer. Students were matched on IQ, age, and sex and then assigned to one of the treatment groups. *Ss* were pretested on a criterion test, presented the material in 17 60-minute classes, and then tested with the same criterion test immediately after instruction. Nichols found that the dependence instruction and structured search instruction were equally effective in teaching plane geometry.

These studies all purported to compare a discovery method of learning with some type of expository learning yet they have yielded inconsistent and often confusing results. Many of these inconsistencies can be attributed to methodological differences among the studies. The most serious problem is the difficulty of replicating the studies because of the absence of detailed descriptions of procedure and operational definitions of the treatments. Wittrock (1966) stated:

> In many of the empirical studies, the treatments are not operationally defined. They are complex and lengthy sequences of stimuli which often differ from other treatments in any number of ways. There often are no principles described which could be used by another researcher to generate the same types of treatments to replicate the study. The independent variable is not isolated or carefully varied. (P. 43)

The studies by McConnell (1934), Thiele. (1938), and Swenson (1949) all used the same general methodology. However, adherence to the teaching method was not carefully controlled. This alone could account for the inconsistent findings. The possibility of incorrect interpretation of the procedure by teachers is also evident in Anderson's (1949) study. First of all, the teachers were included in the treatment method which more closely approached their established method of teaching. During the study they were given freedom in deciding content and the amount of time spent in different types of activities. A written manual and training sessions were the only attempts to standardize the procedure. Anderson may have been studying an interaction of teaching habits and interpretations of procedure rather than two clearly defined methods of presentation.

Worthen (1968) exercised control over the administration of his two treatments by using printed materials as well as teachers' verbalizations. He used the same number and types of examples and equated the degree of verbal presentation for the two methods. Each teacher taught two classes—one in each method—and was provided with detailed descriptions of teaching behavior. Observers' and pupils' ratings were used to ascertain if the teachers

varied their behavior sufficiently to provide a valid test of the two teaching methods.

Ray (1961) and Nichols (1956) also used both teachers' verbalizations and written material. However, the expository method involved mainly teachers' verbalizations while the discovery method involved both verbalizations and written materials. In addition, *Ss* in the expository group learned through group activities while *Ss* in the discovery group learned independently. Thus, the effects of oral versus written and of group versus individual learning activity are confounded with the effects of expository versus discovery presentation.

These comments point up the need for a learning task which allows a more definitive description of the instructional methods being compared. This task should allow the presentation of the same material by different methods and minimize the effect of the teacher. Scott and Frayer (1970) have outlined a standardized learning task which meets these requirements. The main characteristics of this standardized task are that the same amount of information is presented in each treatment, the same examples are used, the same mode of presentation (either oral or written or both) is used, and both treatments are presented to individuals or to a group.

An example of such a task was that used by Scott (1970) in comparing discovery and expository methods of teaching geometry concepts. Sixth-grade students studied prepared lessons during class periods on four consecutive days. Under the expository method the name of the concept was given, followed by positive and negative examples, each accompanied by explicit statements of the relevant attributes of the concept. Under the discovery method, a series of examples was presented. *Ss* were asked to describe each example and to state how examples were alike and how they were different. After all examples were presented, the name of the concept and its relevant attributes were given. The concepts taught and examples used under both methods were identical. Since all lessons consisted of printed material, the effect of variability in instruction by teachers was eliminated.

Scott was interested in both immediate acquisition and retention. He administered a test to some of the students immediately after completion of the lessons and found that there was no significant difference between the discovery and expository groups. Other students were given the same test 1, 11, or 21 days after completion of the lessons. A significant interaction between method and retention interval was noted. The scores of students who learned by the discovery method actually increased over time, while the scores of students who learned by the expository method decreased over time. No tests of significance between mean performances of the groups at the three retention intervals were made. However, the expository group had higher mean scores at 1 day and 11 days, while the discovery group had a higher mean score at 21 days. This result would lend support to some of the claims of proponents of the discovery method. It is desirable to see if this result can be replicated using the same materials and similar methodology.

The present study used the same lessons as Scott, with a few minor modifications. The most significant modification was that six positive and three negative examples were used to teach each concept in the present study, while Scott used eight positive and four negative examples. Also, retention was measured only at 1 day and 21 days after the lessons, deleting the 11-day retention group used by Scott.

II
Method

The experiment was designed to investigate the effects on immediate acquisition and retention of two methods of presenting selected geometry concepts to intermediate grade children. The two methods of presentation were expository and discovery. Assessment of the effect of method of presentation on immediate acquisition of the selected concepts was measured immediately following the lessons; the effect on retention was measured 1 day and 21 days after completion of the lessons.

The following specific questions were asked:

1. Does the level of immediate acquisition of geometry concepts differ for students who are presented the concepts in an expository mode and students who are presented the concepts in a discovery mode?
2. Does the level of retention of geometry concepts differ for students who are presented the concepts in an expository mode and students who are presented the concepts in a discovery mode?

Two pilot studies were run. Pilot Study I had three purposes. The first purpose was to determine the grade level of subjects to be used in the main study. Scott (1970) used sixth-grade students in his study. Sixth graders found the material quite difficult, however, achieving a mean score of only 13.1 out of a possible 28. Pilot Study I was carried out to determine whether the lessons and test were of appropriate difficulty for seventh-grade students. A second purpose for running Pilot Study I was to evaluate the materials and to determine any necessary revisions. A third purpose was to obtain an estimate of the time needed to complete each part of the study. Pilot Study II was run to disclose any problems in the experimental procedures and to determine whether degree of original learning was equivalent under the expository and discovery methods of presentation.

Pilot Study I

Subjects

*S*s were six seventh-grade students. The subjects were randomly selected from students enrolled in remedial reading and/or remedial math classes at Oregon Junior High School in Oregon, Wisconsin. Remedial students were selected since it was thought that this would provide an upper bound for time limits and would be more likely to reveal any problems requiring revision of the materials.

Materials

The materials consisted of two introductory lessons; two quadrilateral lessons each written in two styles, discovery and expository; four placebo lessons; and one test.

Introductory lesson 1. This lesson introduced the concepts ***point***, ***line segment***, and ***line***.

Introductory lesson 2. This lesson introduced the concepts ***closed curve***, ***simple curve***, ***plane***, ***polygon***, ***parallel***, ***adjacent***, ***opposite***, and ***equal length***.

The two introductory lessons provided background information necessary for understanding the quadrilateral concepts. In these lessons a format was used which required the student to respond to questions regarding the concepts. Immediate feedback was provided for these questions.

Quadrilateral lessons. Six examples, four positive and two negative, in the sequence +, −, +, −, +, + were given for each of the four

concepts *quadrilateral*, *rhombus*, *trapezoid*, and *parallelogram*. After two concepts were presented in the above fashion, the first concept was presented again using six different examples in the same sequence, +, -, +, -, +, +. The concepts were presented in this manner and sequence in both the expository and discovery methods. The differences between the lessons occurred in the sequence of definition and examples for each concept and in the statements accompanying each example. In the expository lessons, the definition preceded the examples; in the discovery lessons the definition followed the examples. In the expository lesson each example was accompanied by a statement such as: "Look at this figure. Note that side AB is equal to AC: $\overline{AC}$ is 1"; $\overline{AB}$ is 1"." In the discovery lesson each example was accompanied by a statement such as: "Look at this figure. Measure the side AB. Measure the side AC. What do you find?"

The quadrilateral lessons were presented in prepared booklets similar to those used in the introductory lessons, except for differences in use of questions and feedback. In the quadrilateral lessons, the expository method of presentation included no questions. The discovery method of presentation included questions but provided no feedback during the lesson. Feedback, however, was provided for two questions at the end of the presentation of each concept for the discovery method of presentation. These two questions asked the student to tell which examples were alike and how they were alike. The feedback indicated which examples were alike and gave a definition of the concept in terms of its relevant attributes. The concepts *quadrilateral* and *rhombus* were presented in the first quadrilateral lesson; *trapezoid* and *parallelogram* in the second quadrilateral lesson.

Placebo lessons.

1. *Subtraction.* This lesson presented subtraction as the inverse of addition and showed how addition can be used to check subtraction.
2. *Properties of Addition and Multiplication.* This lesson presented the commutative and associative properties and the identity element in addition and multiplication.
3. *Numeration.* This lesson introduced the numeration system in base ten for numbers having up to seven digits.
4. *Roman Numerals.* This lesson explained the formation of Roman numerals for the numbers 1 through 1100.

Tests of geometry knowledge. A 28-item multiple-choice test developed by Scott (1970) was used in this study. The items of this test required recognition of attribute examples, concept examples and non-examples, relevant attributes, and the concept definition for the concepts *quadrilateral*, *rhombus*, *trapezoid*, and *parallelogram*. Eight completion items were also administered. These items required the student to draw an example of each of the four concepts on graph paper and to list the relevant attributes of each concept. Form BN-1 contained the 18 (14 multiple-choice and 4 completion) items dealing with the concepts of *quadrilateral* and *rhombus*; Form BN-2, the 18 items dealing with the concepts of *trapezoid* and *parallelogram*.

Procedure

The six students were randomly assigned to one of three groups, the expository group (Group E), the discovery group (Group D), or the control group (Group C). The experimenter (*E*) met with each student individually. The student was asked to read the materials aloud and to answer any questions included in the lessons in the presence of *E*. *E* noted any difficulties that the student experienced in reading the materials or in understanding what was being asked of him. Group E received Introductory Lessons 1 and 2, Quadrilateral Lessons 1 and 2 written in the expository mode, and Tests of Geometry Knowledge: Forms BN-1 and BN-2. Group D received Introductory Lessons 1 and 2, Quadrilateral Lessons 1 and 2 written in the discovery mode, and Tests of Geometry Knowledge: Forms BN-1 and BN-2. Group C received Placebo Lessons 1-4. Table 1 outlines the schedule of lessons and tests for each of the three groups.

Results

No statistical analysis was performed on the results of Pilot Study I. The time required for each subject to complete each lesson and test was recorded. Word recognition errors and misinterpretations of questions were noted as a basis for lesson revision. Also, scores on the tests were used to estimate the difficulty of the lessons for seventh graders. These data served as the basis for the following decisions:

1. Seventh-grade students were to be used in Pilot Study II and the main study. The lessons seemed to be of appropriate difficulty for seventh-grade students. The

Table 1. Schedule for Pilot Study I

	Day			
	1	2	3	4
Group E (Expository)	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (Expository) and Test BN-1	Quadrilateral Lesson 2 (Expository) and Test BN-2
Group D (Discovery)	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (Discovery) and Test BN-1	Quadrilateral Lesson 2 (Discovery) and Test BN-2
Group C (Control)	Placebo Lesson 1	Placebo Lesson 2	Placebo Lesson 3	Placebo Lesson 4

test scores ranged from 9 to 20 on a 28-item test for students who studied the geometry lessons and whose grade equivalents on the Paragraph Meaning subtest of the Stanford Achievement Test ranged from 2.9 to 7.7.

2. Introductory Lesson 2 was shortened to insure that it could be completed within a single class period. The material omitted was not needed to successfully complete the quadrilateral lessons.
3. The placebo lessons were shortened and redesigned so that the *S*s were given immediate feedback in a modified linear programming format similar to that used in the Introductory Lessons received by Groups D and E.

Pilot Study II

Pilot Study II compared immediate acquisition under discovery and expository methods of presentation.

Subjects

The initial sample consisted of 79 seventh-grade students; 26 were lost due to absences or failure to complete a lesson, so that results of the study were based on 53 *S*s. The subjects constituted the student population of three mathematics classes at a junior high school in Beloit, Wisconsin. The three classes were selected from a total of 13 classes to represent high, medium, and low achievement levels. A questionnaire (Appendix A) completed by the teachers of these three classes revealed that the students had not previously been taught the concepts *quadrilateral*, *rhombus*, *trapezoid*, and *parallelogram*.

Materials

Introductory Lesson 1, Introductory Lesson 2, Quadrilateral Lesson 1 (discovery), Quadrilateral Lesson 1 (expository), Quadrilateral Lesson 2 (discovery), Quadrilateral Lesson 2 (expository), and Tests of Geometry Knowledge: Forms BN-1 and BN-2 were used in Pilot Study II. These lessons were identical to those used in Pilot Study I, except that Introductory Lesson 2 was shortened, and minor revisions were made in all lessons to simplify wording or clarify ambiguities.

Procedure

*S*s were randomly assigned within class to one of two groups, expository (E) or discovery (D). Both groups were presented Introductory Lessons 1 and 2 on the first two days. On the third day Group E was given Quadrilateral Lesson 1 written in the expository style and Tests of Geometry Knowledge: Form BN-1; Group D was given Quadrilateral Lesson 1 written in the discovery style and Tests of Geometry Knowledge: Form BN-1. On the fourth day Group E was given Quadrilateral Lesson 2 written in the expository style and Tests of Geometry Knowledge: Form BN-2; Group D was given Quadrilateral Lesson 2 written in the discovery style and Tests of Geometry Knowledge: Form BN-2. Table 2 presents the schedule of lessons and tests for Pilot Study II.

Design

The design was a 3 x 2 randomized block design with three classes differing in mean level of achievement (high, middle, and low mathematics achievement) and two types of

Table 2. Schedule for Pilot Study II

	Day			
	1	2	3	4
Group E (Expository)	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (Expository) and Test BN-1	Quadrilateral Lesson 2 (Expository) and Test BN-2
Group D (Discovery)	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (Discovery) and Test BN-1	Quadrilateral Lesson 2 (Discovery) and Test BN-2

quadrilateral lessons (discovery and expository). Subjects within each class were randomly assigned to one of the two treatments.

Results

A more complete summary of the results from Pilot Study II will be presented in Chapter III. The following descriptive results are presented to serve as background for changes made in the procedure for the main study:

1. Some members of the discovery group were unable to complete the lesson and test during class periods on the third and fourth days of the experiment. The following changes were made to reduce the time required.
   a. The test was shortened by eliminating the completion questions. An item analysis revealed that they did not discriminate as well as the multiple choice items, while requiring relatively more time.
   b. The students were allowed to circle their answers directly on the test booklet rather than using an answer sheet. This reduced the time needed for instructions as well as the time needed to complete the test.
   c. The two quadrilateral lessons were shortened by reducing the total number of examples for each concept from 12 to 9. Five examples were used to present the concept the first time. The sequence of positive and negative examples was +, −, +, −, +. The second presentation used four different examples in the sequence +, −, +, +. The sequences of positive and negative examples were the same for all four concepts.
2. Many students were uncertain how to measure lines with a ruler. Therefore, a short lesson on the use of the ruler and measuring exercises were added to the first day's lesson.

Main Study

Subjects

The initial sample consisted of 318 seventh-grade students in 11 classes. These students constituted the entire seventh-grade student population of a junior high school in Beloit, Wisconsin. This was a different school than the one used in Pilot Study II. Ninety *S*s were lost because of absences, so that the results were based on 228 *S*s. A questionnaire (Appendix A) completed by the students' mathematics teachers revealed that approximately 50 percent of the students had been introduced to the concept *quadrilateral* and 10 percent of the students had been introduced to the concept *parallelogram*. The concepts *trapezoid* and *rhombus* had not been introduced to the students prior to the beginning of the study.

Materials

Introductory Lesson 1, Introductory Lesson 2, Quadrilateral Lesson 1 (discovery), Quadrilateral Lesson 1 (expository), Quadrilateral Lesson 2 (discovery), Quadrilateral Lesson 2 (expository), Placebo Lessons 1-4 with feedback for each question, and Tests of Geometry Knowledge: Forms BN-1 and BN-2 were used for this study. In addition, Form BN of Tests of Geometry Knowledge, comprised of all items from Forms BN-1 and BN-2, was used.

Procedure

*S*s were rank ordered on the basis of their scores on the Applied Arithmetic subtest of the Stanford Achievement Test. The first eight *S*s

Table 3. Schedule for the Main Study

	Group	Day 1	2	3	4	5	25
Discovery	1	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (D) Test BN-1	Quadrilateral Lesson 2 (D) Test BN-2		
	2	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (D)	Quadrilateral Lesson 2 (D)	Test BN	Test BN
	3	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (D)	Quadrilateral Lesson 2 (D)		Test BN
Expository	4	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (E) Test BN-1	Quadrilateral Lesson 2 (E) Test BN-2		
	5	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (E)	Quadrilateral Lesson 2 (E)	Test BN	Test BN
	6	Introductory Lesson 1	Introductory Lesson 2	Quadrilateral Lesson 1 (E)	Quadrilateral Lesson 2 (E)		Test BN
Control	7	Placebo Lesson 1	Placebo Lesson 2	Placebo Lesson 3	Placebo Lesson 4	Test BN	Test BN
	8	Placebo Lesson 1	Placebo Lesson 2	Placebo Lesson 3	Placebo Lesson 4		Test BN

in the ranking were randomly assigned to eight treatment groups. The second eight *S*s were then randomly assigned and so on until all *S*s had been assigned to treatment groups.

A series of four lessons was presented to the *S*s over a period of four days. Table 3 presents the schedule of lessons and tests for each of the eight treatment groups. On Days 1 and 2, Groups 1-6 studied Introductory Lessons 1 and 2. On Days 3 and 4, Groups 1-3 were presented Quadrilateral Lessons 1 and 2 written in the discovery mode.

A test to determine level of concept attainment was administered to *S*s in Group 1, the immediate acquisition group, immediately following completion of the lessons. Since the quadrilateral concepts were taught in two lessons, the immediate acquisition group received after each lesson a test (Form BN-1 or BN-2) dealing only with the concepts included in that lesson. There was no counterbalancing of the order of concept presentation since the attainment of concepts presented on the fourth day was contingent upon the completion of the lesson presented on the third day. Group 2 was administered the test 1 day after the completion of the lessons. The test was readministered to Group 2 21 days after completion of the lessons. Group 3 was administered the test 21 days after the completion of the lessons.

Groups 4-6 corresponded to Groups 1-3. The only difference was that the Quadrilateral Lessons 1 and 2 were written in the expository mode rather than the discovery mode. Groups 7 and 8 received placebo lessons in place of the introductory and quadrilateral lessons. They were, however, given the same tests as Groups 1-6. Subjects in Group 7 took the test 1 day and 21 days after completion of the placebo lessons, and Group 8, 21 days after the lessons.

All materials were prepackaged by day and class with *S*s' names on the lesson booklets to insure that each *S* received the proper lesson. On Day 1 the proctor distributed pencils, rulers, cardboard strips (used to cover the answers listed in the right-hand column of lessons with immediate feedback), and booklets. Instructions (Appendix B) concerning the procedure to be followed in completing the lessons were read to the students. Difficult words, listed on the first page of the booklets, were read to the *S*s and any questions concerning procedure were answered. Each student recorded the starting time and

Table 4. Number of Subjects in Each Treatment by Achievement Group

Arithmetic Achievement Level	Group							
	Discovery			Expository			Control	
	Immediate	1 Day	21 Days	Immediate	1 Day	21 Days	1 Day	21 Days
High	7 (3)	8 (2)	6 (4)	10 (0)	7 (3)	7 (3)	8 (2)	8 (2)
Medium	15 (2)	11 (6)	16 (1)	11 (6)	12 (5)	16 (1)	13 (4)	11 (6)
Low	4 (7)	11 (0)	7 (4)	10 (1)	8 (3)	8 (3)	7 (4)	7 (4)
Total	26	30	29	31	27	31	28	26

Note: The number of subjects lost from each group due to absence is given in parentheses.

studied his lesson on an individual basis. Upon completion of his lesson the student recorded his finishing time and worked on an assignment given by the classroom teacher. After all students had finished, a lesson on the use of a ruler was given, followed by an exercise requiring measurement of lines using a ruler. All materials were collected at the end of the class period.

On Days 2, 3, and 4 the same general procedure was followed. On Day 2 after all students had completed their lessons they were given instructions for the test which some students would take on Day 3. On Days 3 and 4 students indicated completion of their lesson by raising their hand. Each student in the immediate acquisition group was given a test upon completion of his lesson.

On Day 5 those students in the 1-day retention group were given a test. Students not in the 1-day retention group worked on an assignment given by the classroom teacher. On Day 25 all students received a test. Two proctors, substitute teachers in the Beloit school system, were hired to administer the lessons and tests. Checks by two independent observers revealed that both proctors were adhering to the general procedures.

Design

The experimental design was a 3 x 3 x 3 incomplete randomized block design with three levels of arithmetic achievement (high, medium, and low), three types of lessons (discovery, expository, and control), and three retention intervals (immediate, 1 day, and 21 days). Achievement levels were formed by combining the blocks used for assignment of subjects to treatment groups. Criteria for combining the blocks will be described in Chapter III. To reduce the number of groups, the immediate acquisition control condition was deleted from the design. The resulting design had 24 groups, three levels of arithmetic achievement with eight treatment groups at each level. The number of subjects in each cell of the design is shown in Table 4.

Subjects in the 1-day retention groups were retested at 21 days. This permitted a repeated measures analysis as well as an independent groups analysis of the effects of discovery and expository learning on retention.

III
Results

Pilot Study II

Four dependent measures, a score for the multiple choice items on Form BN-1, a score for the production items on Form BN-1, a score for the multiple choice items on Form BN-2, and a score for the production items on Form BN-2, were obtained for each *S*. These four scores made it possible to detect differences in the difficulty of the two forms and in the sensitivity of the two types of questions. Table 5 presents the number of subjects and the means and standard deviations of multiple choice, production, and total scores for treatment groups by class.

Table 5. Number of Subjects and Observed Means and Standard Deviations of Multiple Choice, Production, and Total Scores for Treatment Groups by Class in Pilot Study II

Ability Level of Class	Treatment Group: Discovery Immediate	Treatment Group: Expository Immediate	M
High			
Multiple Choice[a]	26.13 (3.40)	26.58 (1.44)	26.36
Production[b]	7.37 (.74)	7.33 (1.38)	7.35
Total	33.50	33.91	33.71
	N = 8	N = 12	N = 20
Middle			
Multiple Choice	13.50 (3.40)	19.42 (5.62)	16.46
Production	2.33 (1.76)	4.58 (2.40)	3.46
Total	15.83	24.00	19.92
	N = 6	N = 12	N = 18
Low			
Multiple Choice	9.75 (3.30)	16.36 (5.82)	13.06
Production	1.75 (.50)	4.27 (2.50)	3.01
Total	11.50	20.63	11.98
	N = 4	N = 11	N = 15
M			
Multiple Choice	16.46	18.06	17.26
Production	3.82	5.39	4.61
Total	20.28	23.45	21.87
	N = 18	N = 35	N = 53

Note: Standard deviations are given in parentheses.

[a]There were 28 multiple choice items.
[b]There were 8 production items.

Table 6. Analysis of Variance of Subtest Scores on Forms BN-1 and BN-2 in Pilot Study II

Source	*df*	*MS*	*F*	Probability
Total—Between Subjects	53			
Mean	1			
Between Cells	5			
Class	2	77.30	36.44	< .0001*
Treatment	1	20.13	9.49	< .0035*
Class x Treatment	1	5.95	2.81	< .0707
Between Subjects/Within Cells	47	2.12		
Total—Within Subjects	159			
Form of Test	1	14.53	19.72	< .0001*
Form x Cells	5			
Form x Class	2	.97	1.31	< .28
Form x Treatment	1	1.65	2.24	< .14
Form x Class x Treatment	2	.35	.47	< .63
Form x Subjects/Within Cells	47	.74		
Type of Item	1	737.83	1120.55	< .0001*
Type x Cells	5			
Type x Class	2	18.83	28.59	< .0001*
Type x Treatment	1	4.50	6.83	< .012*
Type x Class x Treatment	2	1.00	1.52	< .23
Type x Subjects/Within Cells	47	.66		
Type x Form	1	9.77	26.66	< .0001*
Type x Form x Cells	5			
Type x Form x Class	2	.28	.76	< .48
Type x Form x Treatment	1	.07	.19	< .67
Type x Form x Treatment x Class	2	.17	.45	< .64
Type x Form x Subjects/Within Cells	47	.37		

*Significant at or beyond .05 level chosen.

A repeated measures analysis of variance with a 2 x 2 within subjects design (two forms of test and two types of item) was performed on the data in Table 5. A summary of this is presented in Table 6. Scores obtained under the expository method of presentation were significantly higher than those obtained under the discovery method of presentation. This effect remained significant after achievement (class) differences were removed. The class effect was significant, but there was no treatment by class interaction.

Multiple choice items were found to be more sensitive than production items regardless of form. This led to the elimination of the production items on the test used for the main study.

Scores on the Arithmetic Application and Paragraph Meaning subtests of the Stanford Achievement Test were available for each *S*. Correlations between the total test score for Forms BN-1 and BN-2 combined and the Arithmetic Application and Paragraph Meaning scores were calculated to determine which score would be used as a blocking factor in the main study. These correlations were .47 and .34, respectively. Since the correlation with Arithmetic Application was higher than that with Paragraph Meaning, the Arithmetic Application score was used as a blocking variable in the main study.

Main Study

Scores on the Arithmetic Application subtest of the Stanford Achievement Test were obtained from school records as a basis for stratifying subjects. One dependent measure, the total score on the 28-item multiple choice test, was obtained for each *S*. *S*s in the one-day retention groups had an additional score; Test BN was readministered to these *S*s after the 21-day retention interval. Scores for each student on the Arithmetic Application test and on Test BN are given in Appendix C.

Psychometric Characteristics of Test BN

An item analysis (Baker & Martin, 1968) was performed on the data for all subjects. In general, the items met the criteria for "good" items established by Harris (1968). Twenty-four of the 28 items had item difficulties between .30–.70. The point biserial correlations for the correct alternatives were generally well above +.50 and for the incorrect alternatives generally well below -.20. The Hoyt reliability estimate for the 28-item test was .90.

Analysis of the Data

Independent groups analysis. The blocks used for assigning *S*s to treatment groups were combined to form three achievement levels. Each of the three achievement blocks which resulted represented a unique range of Arithmetic Application scores, and all subjects in a block used for random assignment of *S*s were assigned to the same achievement block. The three achievement levels combined with the eight treatment groups to yield 24 treatment by achievement groups. The number of subjects, means, and standard deviations for each of the treatment by achievement conditions are given in Table 7.

An analysis of variance using Finn's (1968) multivariance computer program was performed on the data in Table 7. Since the number of *S*s in the cells varied, the design was non-orthogonal and the effects were not independent. In general, the major groupings were tested in the order indicated. Within each group all effects were removed last.

The results of this analysis are shown in Table 8. A significant difference between discovery and expository modes of presentation on immediate acquisition was found.

Table 7. Number of Subjects and Observed Means and Standard Deviations of Scores for Treatment Groups by Achievement Level in the Main Study

Arithmetic Achievement	Treatment Group								
	Discovery			Expository			Control		
	Immediate	1 Day	21 Days	Immediate	1 Day	21 Days	1 Day	21 Days	M
High	23.43 (5.26) N = 7	23.11 (5.49) N = 8	17.83 (6.01) N = 6	27.00 (1.49) N = 10	25.00 (5.00) N = 7	21.00 (4.80) N = 7	14.75 (4.59) N = 8	17.89 (5.99) N = 8	21.25 N = 61
Medium	18.20 (5.97) N = 15	13.36 (4.86) N = 11	13.88 (4.46) N = 16	21.63 (6.00) N = 11	16.75 (6.38) N = 12	14.63 (5.56) N = 16	11.27 (4.45) N = 13	9.73 (3.85) N = 11	14.93 N = 105
Low	12.75 (2.22) N = 4	9.36 (3.20) N = 11	9.57 (1.99) N = 7	17.50 (6.95) N = 10	14.50 (7.46) N = 8	10.25 (3.96) N = 8	8.86 (4.22) N = 7	8.00 (2.52) N = 7	11.35 N = 62
M[a]	18.13 N = 26	15.28 N = 30	13.76 N = 29	22.04 N = 31	18.75 N = 27	15.29 N = 31	11.63 N = 28	11.87 N = 26	15.84 N = 228

Note: Standard deviations are given in parentheses.

[a]Unweighted mean for high, medium, and low achievement groups.

Table 8. Analysis of Variance of Total Scores on Test BN in the Main Study

Source	*df*	*MS*	*F*	Probability
Mean				
Between Achievement Blocks	2	1621.46	63.60	< 0.0001*
Time				
Immediate Recall vs Recall during Retention Interval	1	1738.28	68.18	< 0.0001*
1 Day Retention vs 21 Day	1	44.34	1.74	< 0.19
Achievement Blocks x Time	4	5.30	0.21	< 0.93
Treatments				
D vs E on Immediate Recall	1	105.81	4.15	< 0.04*
D vs E during Retention given the Immediate Recall Differences	1	1.71	0.07	< 0.80
Control vs average of D + E during Retention	1	663.62	26.03	< 0.0001*
Treatment x Time during Retention				
D vs E at 1 Day vs D vs E at 21 Days	1	32.32	1.27	< 0.26
Control vs average of D + E at 1 Day vs 21 Days	1	49.17	1.93	< 0.17
Achievement Blocks x Treatment & Achievement Blocks x Treatment x Time	10	34.34	1.35	< 0.25
Between Subjects within Cells (error)	208	25.50		

*Significant at or beyond the .05 level chosen.

After removing this difference between groups on immediate acquisition, no difference between discovery and expository groups during the retention interval was noted. A comparison between the average of the discovery and expository group scores (adjusted for differences in immediate acquisition) and the control group scores revealed that the learning materials were effective, since the performances of the discovery and expository groups were superior to those of the control groups. during the retention interval. Figure 1 shows the mean total score under each treatment as a function of time. Means for both the discovery and expository groups decreased over time. Both discovery and expository groups performed better at each retention interval than did the control group.

In Figure 1 it appears that there may have been a difference in the relative amount of forgetting following the two treatments, i.e., the mean scores for the expository group decrease more than the mean scores for the discovery group. However, the interaction of D versus E and time was not statistically significant. The difference between levels of

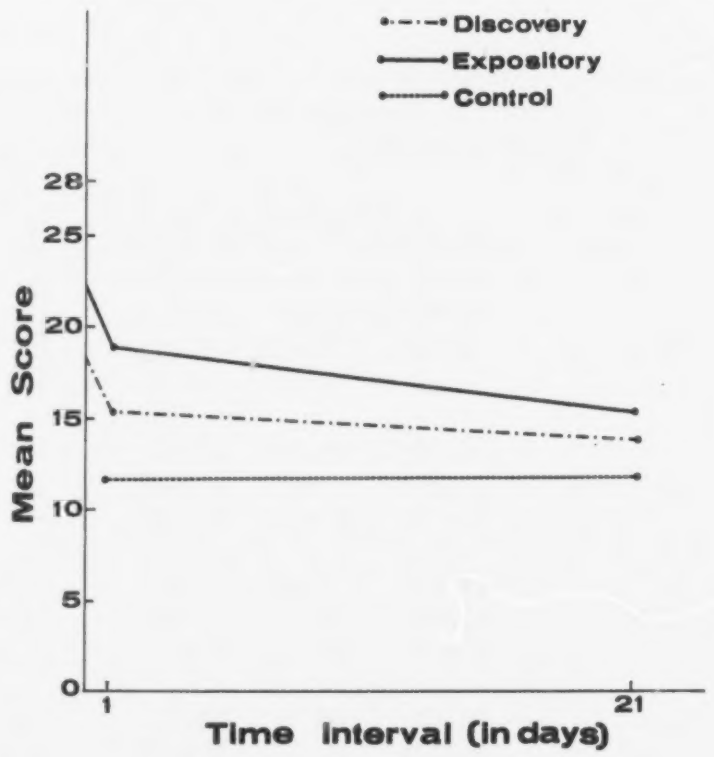


Figure 1. Mean scores of discovery, expository, and control groups immediately, 1 day, and 21 days after completion of lessons (independent groups analysis).

Table 9. Mean Scores and Changes in Scores for Repeated Administrations of the Same Test to Discovery, Expository, and Control Groups at 1 Day and 21 Days

	Treatment		
	D	E	C
1 day	14.71	18.22	11.53
21 days	15.03	15.81	12.13
Change (1 day to 21 days)	+0.32	-2.41	+0.60

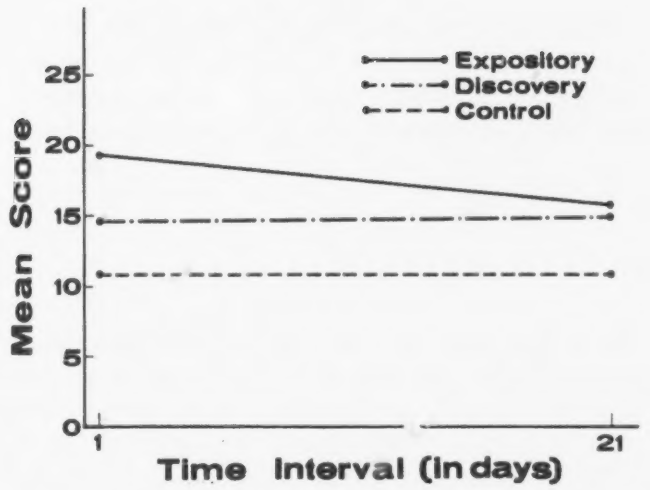


Figure 2. Mean scores of discovery, expository, and control groups 1 day and 21 days after completion of lessons (repeated measures analysis).

retention for D and E is estimated to decrease by 2.05 score units from 1 day to 21 days. But the standard error of this estimate is 1.82 so that even a 90 percent confidence interval centered on the estimate would include 0.

A secondary question of interest in this study was "Do children at one of the achievement levels benefit more from either the discovery or expository method than children at the other achievement levels?" Although there was a significant achievement effect, neither the achievement by treatment nor the achievement by treatment by time interaction was significant. Thus, the evidence suggests that the relative effectiveness of discovery and expository learning does not vary with the achievement level of the student.

Repeated measures analysis. The analyses reported earlier in this chapter involved differences between the scores of independent groups who varied in terms of treatment and the time interval between completion of the lesson and administration of the test. A second way of examining the effects of discovery and expository learning on retention entailed the use of a repeated measures design. Students in the 1-day retention groups were administered the test one day after completion of the lesson and were readministered the test 21 days after completion.

An analysis of variance was performed on the changes in score between the two testings. The means for the two testings and the change scores are presented in Table 9. Figure 2 shows the mean total score for each treatment group at 1 day and at 21 days. The analysis of variance on these change scores is summarized in Table 10. The result of major interest is that there is a significant difference in change scores among the three groups. The expository group differed significantly from both the discovery group and the control group. The expository group decreased in performance from 1 day to 21 days, while the discovery and control groups did not decrease.

It should be noted that the analysis of change scores in the repeated measures design corresponds to the treatment by time interaction in the independent groups design. The change scores in the repeated measures analysis differed significantly as a function of treatment, however, while the treatment by time interaction in the independent groups analysis was not significant. The difference in results obtained by the two methods of analysis can be attributed in part to the relative sensitivity of the analyses. The estimate of change is about the same for the independent groups design as for the repeated measures design. However, the standard error for the repeated measures approach is only 50 to 60 percent as large as the standard error for the independent groups design.

The difference in the results of the repeated measures and independent group analyses cannot be attributed to the effects of repeated testing. The 21-day scores of the "1-day group" were compared with the 21-day scores of the "21-day group." No group differences or group by treatment interactions were found, suggesting that testing at 1 day did not markedly affect the scores of the "1-day group" at 21 days.

The results of the repeated measures analysis correspond more closely to Scott's results than do the results from the independent groups analysis. In Table 9 it can be seen that the mean scores of the discovery group actually increase over time from 14.71 to 15.03 while the scores of the expository group de-

crease from 18.22 to 15.81. However, it should be noted that even with the increase in the performance of the discovery group and the decrease in the performance of the expository group, the mean score of the expository group is still slightly higher than that of the discovery group.

Table 10. Analysis of Variance of Change Scores in the Repeated Measures Design

Source	*df*	*MS*	*F*	Probability
Mean Change (after allowing for Treatment x Time)	1	23.21	1.70	< 0.20
Change Comparisons[a]	2	79.09	5.79	< 0.004*
D vs E[b]	1	112.91	8.26	< 0.005*
Control vs average of D + E[b]	1	56.46	4.13	< 0.045*
Lack of Fit[c] (Achievement Blocks x Time and Treatments x Achievement Blocks x Time)	8	11.40	0.83	< 0.58
Subjects x Time Within Cells	3	13.67		

*Significant at or beyond the .05 level chosen.

[a]Assuming no lack of fit (as shown above, the data do not show significant lack of fit).
[b]Each 1 *df* test was done after *SS* for the other 1 *df* source had been removed.
[c]One cell is missing.

IV
Discussion and Conclusions

Discussion

This study was performed to determine the differential effects on immediate acquisition and retention of two methods of presenting geometry concepts to seventh-grade students. The results of this study indicated that the written lessons were effective in teaching geometry regardless of the mode in which they were written. *Ss* in the two experimental groups performed better on a test of geometry knowledge than did *Ss* in the control group. Method of presentation differentially affected immediate acquisition, with *Ss* in the expository group demonstrating superior performance on a test given immediately after the presentation of the lessons. The evidence concerning the effect of discovery and expository methods of presentation on retention is less clear-cut. An independent groups analysis revealed no significant difference between treatment groups during a 21-day retention interval after the difference in immediate acquisition was removed. A repeated measures analysis, on the other hand, revealed that *Ss* in the discovery groups tended to forget less than *Ss* in the expository groups.

The significant difference between the immediate acquisition levels of the discovery and expository groups leads to difficulty in interpreting the differential retention indicated by the repeated measures analysis. We would like to know the effect of method of presentation on retention, given the same amount of initial learning. Since the amount learned was not equal we must temper our finding of differential retention with the qualification of different levels of initial learning. Related to this is the fact that even though the discovery group forgot less than the expository group, the lower initial score of the discovery group resulted in the two groups being roughly equal at the end of the retention interval.

The methodology and materials used in this study were very similar to those used by Scott (1970). Scott found no significant difference between discovery and expository groups on immediate acquisition, but a significant difference in retention over the 21-day interval. The scores of the discovery group increased, while the scores of the expository group decreased. The results of the repeated measures design offer some support for Scott's observation that the discovery group had better long-term retention. However, since there is a difference between the results of the two studies on immediate acquisition and since only one of the analyses supports Scott's findings concerning retention, it would be well to compare these two studies.

The lessons used in the present study were basically the same as those used by Scott with one major difference. Scott used a total of eight positive and four negative examples for each concept. Because of a limit in the class time available in the present study, only six positive and three negative examples were used. The ratio of positive to negative examples remained the same, but the number of examples was reduced. The number of examples was the same for both treatment groups. There may have been an interaction between the number of examples and the treatments which could have, in part, accounted for the differences in results. However, this seems unlikely. The method of presenting concepts through negative as well as positive examples is not widely used. Not enough is known about the method to allow the specification of the optimum number of positive and negative examples or to suggest that reducing the number of examples would cause the contradictory results.

The methodologies of these two studies

were similar and followed the guidelines established by Scott and Frayer (1970). However, the age of the subjects and the number of subjects per treatment group differed for the two studies. Scott used sixth-grade students, but noted that students at this grade level found the material difficult. On the basis of this finding, seventh-grade students were used in the present study.

It is reasonable to ask if the material was perhaps too easy for seventh-grade students; i.e., if a ceiling effect may have affected results. Suggestion of a ceiling effect was found for only one of the 24 treatment by achievement groups. The high achievement-immediate acquisition group taught under the expository mode had a mean score of 27.0 out of a possible 28 with a standard deviation of 1.49. This ceiling effect would have minimized the difference between the two methods. Therefore, the superiority of expository over discovery on initial acquisition may have been even greater if there had not been a ceiling effect. No ceiling effects were noted for the retention groups, however, and this is where such effects might be expected to account for the difference in results between Scott's study and the present study.

As expected, the overall performance of the seventh-grade students in the present study was superior to the overall performance of Scott's sixth-grade students on the same 28-item test. The overall mean in Scott's study was 13.1; in the present study, 15.8. The difference would probably have been larger if the same number of examples had been used to teach concepts in each study.

The one difference between the two studies which lends more credence to the results of the present study than Scott's study lies in the number of subjects in each treatment group. The mean number of *Ss* in each treatment group in Scott's study was 17 and the range was 16-19. In the present study the mean number was 29 and range was 26-31. The increased cell size in the present study may have produced more stable results than those obtained by Scott.

A second finding of this study was that differences in achievement resulted from discovery and expository modes of presentation even when extraneous variables were carefully controlled. Most of the earlier studies comparing discovery and expository modes of presentation relied on the verbalizations of teachers or on a combination of written materials and teachers' verbalizations. It was often difficult to assure strict adherence to prescribed methods within an experiment or to replicate methods between two experiments when the lessons were presented verbally by teachers. The materials used in the present study were entirely in written form for both treatment groups. The same examples were used for both methods of presentation. Also, the materials were equated in terms of the amount and type of information contained in each type of lesson.

Using written materials for both discovery and expository treatments and equating the amount and type of information given in the lessons eliminated many of the criticisms of former studies. However, there was one variable which was not equated between treatments—the time spent studying each lesson. Each student was allowed to proceed at his own pace and was allowed to finish the lesson. The time needed to complete the lessons turned out to be an important variable. *Ss* in the discovery group spent about 50 minutes completing the two quadrilateral lessons, while *Ss* in the expository group spent about 15 minutes completing the lessons. It would seem reasonable to assume that the more time spent on the lessons, the more complete the learning. However, this was not true. The expository group, which spent about one-third as much time as the discovery group, showed superior performance on the test of immediate acquisition, and retained as much of the material as did those in the discovery group in the independent measures design. Even though the repeated measures analysis revealed less forgetting by the discovery groups, the means of the two groups were roughly equal after the retention interval.

Conclusions

The main conclusion of the present study is that method of presentation differentially affected immediate acquisition and may have differentially affected retention. The expository method led to superior initial acquisition, but in terms of mean performance after 21 days both methods were equally effective. These results do not support the claims of many who feel that the discovery mode of presentation is superior to the expository mode of presentation. To the contrary, they tend to indicate the superiority of the expository method over the discovery method. *Ss* in the expository groups spent less than one-third as much time studying the lessons as *Ss* in the discovery groups, yet still had superior immediate acquisition scores. It is also significant that they demonstrated performance equal to that of

the discovery group after a three-week retention interval. Using efficiency as a criterion, it would appear that the expository method is superior to the discovery method.

The results of this study should not lead one to conclude that the expository method is superior to the discovery method under all conditions. This study has compared only immediate acquisition and retention of concepts. The validity of many of the claims of advocates of the discovery method were not tested. Additional research is needed to determine whether discovery learning has effects not tested in this study. For example, is discovery learning intrinsically motivating? Does it develop problem solving ability? The specific effects of discovery learning should be delineated to permit the teacher to make a wise choice of teaching method to reach particular objectives.

References

Anderson, G. L. Quantitative thinking as developed under connectionist and field theories of learning. In E. Swenson *et al.*, *Learning theory in school situations*. Minneapolis: University of Minnesota Press, 1949. Pp. 40-73.

Baker, F. B., & Martin, T. J. FORTAP: A FORTRAN test analysis package. Madison: Wisconsin Research and Development Center for Cognitive Learning, 1968.

Finn, J. D. Multivariance—univariate and multivariate analysis of variance and covariance: A FORTRAN IV program. Version 4. Buffalo: Department of Educational Psychology, State University of New York at Buffalo, 1968.

Harris, M. *Some methodological suggestions for construction of an objective measurement instrument.* Wisconsin Research & Development Center for Cognitive Learning, Technical Memo No. M-1968-2 (Revised), 1968.

McConnell, T. R. Discovery vs. authoritative identification in children. *Studies in Education*, 1934, *2* (5), 13-60. Cited in L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal.* Chicago: Rand McNally, 1966.

Morrisett, L. N. Further reflections. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal.* Chicago: Rand McNally, 1966. Pp. 177-180.

Nichols, E. D. Comparison of two approaches to the teaching of plane geometry. Unpublished doctoral dissertation, University of Illinois, 1956.

Ray, W. E. Pupil discovery vs. direct instruction. *Journal of Experimental Education*, 1961, *29*, 271-280.

Scott, J. A. *The effects on short- and long-term retention and on transfer of two methods of presenting selected geometry concepts.* Wisconsin Research & Development Center for Cognitive Learning, Technical Report No. 138, 1970.

Scott, J. A., & Frayer, D. A. *Learning by discovery: A review of the research methodology.* Wisconsin Research & Development Center for Cognitive Learning, Working Paper No. 64, 1970.

Swenson, E. J. Organization and generalization as factors in learning, transfer, and retroactive inhibition. In E. J. Swenson *et al.*, *Learning theory in school situations.* Minneapolis: University of Minnesota Press, 1949. Pp. 9-39.

Thiele, C. L. *The contribution of generalization to the learning of addition facts.* New York: Teachers College, Columbia University, 1938.

Wittrock, M. C. The learning by discovery hypothesis. In L. S. Shulman & E. R. Keislar (Eds.), *Learning by discovery: A critical appraisal.* Chicago: Rand McNally, 1966. Pp. 33-76.

Worthen, B. R. Discovery and expository task presentation in elementary mathematics. *Journal of Educational Psychology*, 1968, *59* (Monogr. Suppl. No. 1, Part 2).

GPO 827-442-4

Appendix A
Teacher Questionnaire: Knowledge of Geometry Concepts

KNOWLEDGE OF GEOMETRY CONCEPTS

1. Prior to the experiment on February 8, 1971, which of the following geometry concepts had your class studied during the current school year?

	YES	NO	If yes, please give approximate date
point			
line segment			
line			
closed curve			
simple curve			
plane			
polygon			
parallel			
adjacent			
opposite			
equal length			
quadrilateral			
rhombus			
trapezoid			
parallelogram			

Teacher _____

Class Period _____

2. Prior to the experiment on February 8, 1971, what degree of mastery do you think the majority of students in your class possessed regarding the following concepts?

	Completely unfamiliar	Some knowledge, but not mastery	Mastery
point			
line segment			
line			
closed curve			
simple curve			
plane			
polygon			
parallel			
adjacent			
opposite			
equal length			
quadrilateral			
rhombus			
trapezoid			
parallelogram			

Teacher _____

Class Period _____

25/26

Appendix B
Instructions to Students

Day 1 Instructions

Good morning (afternoon).

My name is _____. I am working with some educational psychologists at the University of Wisconsin in Madison. These psychologists are trying to find out how to make it easier for students to learn mathematics. They have written some lessons which you will study this week. After you have completed the lessons, you will be given a short test to see how much you learned. Please do the best job you can on both the lessons and the tests. If you do, you will learn some geometry, and more than that you will help psychologists find ways to make learning easier for other students.

(Hand out supplies.)

Each of you will receive a ruler, pencil, cardboard strip, and a lesson with your name on it. Do not open the lesson until I tell you to. Not all of you will receive the same lessons during this week so don't worry if your lesson doesn't look like your neighbor's lesson. All of the lessons are extremely important. The type of lesson you receive is in no way related to how smart you are.

Check to see if your name is spelled correctly. If it isn't, correct it. Write the name of your teacher and the class period on the line which says "teacher." Fill in your grade and the date but don't write anything on the lines which say "starting time" and "finishing time."

(Write teacher's name, the period, and the date on the board.)

This lesson may be different from other lessons you have done. Here is how it works. The pages in your lesson will look like this. (Open to any page.) This side has questions for you to answer. The other side has the correct answers. When you do the lesson you should cover the answers with the piece of cardboard, like this. (Show them.)

After you write your answer to the questions, move the cardboard down just far enough so that you can see if the answer you wrote is correct. If it is, go on to the next question. If you make a mistake and find that the answer you wrote down is not right, just draw a line through it and write the correct answer beside it. By making corrections like this we will know which questions are too hard.

Turn to the page which says "Word List." If you do not have such a page just wait until I tell you to start. (Read the heading, pronounce each word, and have the students repeat it.) Can you pronounce each word?

When I finish with the instructions, you will go through the rest of the lesson by yourselves. If you have any questions or come to any words that you do not know, raise your hand and I will help you. If you are not sure of something you have learned, you may look back at a page you have already done.

When you finish, write the exact time where it says "finishing time." Work quietly on your assignment until everyone is finished. Then I will collect your lessons. Take your time so that you will understand and be able to answer questions later. Are there any questions? Write the exact time it is now where it says "starting time." (Tell them the time and write it on the board.) Begin working.

(When everyone is finished, collect the lessons.)

(Draw a ruler on the board.)

On Wednesday and Thursday some of you will use the plastic ruler. We will review how

to use it so that you won't have any trouble when you need to use it.

(Hand out the ruler worksheets.)

Line up your ruler with the one that is at the top of your sheet. (Demonstrate.) Notice that the edge with the big numbers 1 through 6 is on top. This is the edge that you will use.

These numbers are in a box. (Outline the box.) When you measure, you will begin measuring from the end of the box, not from the end of the ruler.

When you measure you will have to know where 1/4 of an inch, 1/2 of an inch and 1 inch are on the ruler. I will show you on the board. Mark the picture of the ruler the same way as I mark the ruler on the board.

The fourth line from the edge of the box shows you where 1/4 of an inch is. The line which almost touches the bottom of the box shows you where 1/2 of an inch is. This line (point) shows where 3/4 of an inch is and this line that reaches to the bottom of the box shows you where 1 inch is. (Identify 1/4, 1/2, 3/4, 1, 1 1/4, 1 1/2, 1 3/4, 2, 2 1/4, 2 1/2, 2 3/4, and 3 inches.)

(Check that each student has marked his picture correctly. Students only need to be concerned with accuracy of 1/4 of an inch.)

Now answer the questions on the worksheet. We will check your answers when everyone is finished.

(Read the answers or if there is time let students supply the answers. Answer any questions.)

How many of you got all the answers correct? (Note a rough estimate of the percentage.)

(Collect materials. Students may keep the worksheets. Thank the students for their cooperation.)

Day 2 Instructions

Good morning (afternoon).

(Distribute materials.)

Does everyone have a pencil, ruler, cardboard strip, and a lesson with his name on it? Write the name of your teacher and the class period where it says "teacher." Fill in your grade and the date. (Write name, period, and date on board.)

Today's lesson is just like yesterday's. You will use the cardboard strip in the same way and make corrections in the same way as you did yesterday.

Now turn to the page which says "Word List." If you do not have such a page, just wait until I tell you to begin working. (Read the page as before.) Are there any questions?

When you finish, put the time you finish on the front cover. Raise your hand and I will collect your lesson. The pencils, rulers, and strips will be collected when everyone is finished.

Remember, if you have any questions, just raise your hand and I will help you. The starting time is _____. (Write it on the board.) Begin working.

(When everyone is finished, collect the supplies.)

Tomorrow some of you will take a test. So that you will not have any difficulty in taking it, I will read the instructions to you now. (Read the instructions slowly and go through the examples on the board. Answer any questions.)

Day 3 Instructions

Good morning (afternoon).

(Distribute materials.)

Does everyone have a pencil, ruler, cardboard strip, and a lesson with his name on it? Write the name of your teacher and the class period where it says "teacher." Fill in your grade and the date.

(Write information on the board.)

Today's procedure will be a little different than yesterday's. First of all, the lessons are not quite the same. Not all of you who used the cardboard strip Monday and Tuesday will use the cardboard strip today and some of you who do will use it only on a couple of pages. So when you come to a page which has a line down the side like this (show them), use the cardboard strip as you did before.

Another difference is that after you finish your lesson, some of you will be given a short test. If you receive a test be sure to fill in the information on the front cover and to follow the directions.

Now turn to the Word List page.

(Read the page as before.)

Are there any questions?

When you finish with the lesson write the finishing time on the front cover and raise your hand. When I collect your lesson I will give some of you a test. Raise your hand when you finish the test.

Are there any questions? Put the time it is now on the cover where it says "starting time."

(Tell them the time.)

Begin working by reading the next page.

Day 4 Instructions

Good morning (afternoon).

Today you will receive the last lesson. The procedure today will be just like yesterday's.

(Pass out supplies.)

Does everyone have a pencil, cardboard strip, ruler, and a lesson? Check your name. Write the teacher's name and the class period where it says "teacher." Fill in the date and grade.

(Write this information on the board.)

Turn to the Word List.

(Read the page as before.)

Are there any questions?

When you are finished with the lesson, write the finishing time on the cover. Raise your hand and I will collect your lesson and give some of you a test. If you receive a test, be sure to fill in the information on the cover and to follow the directions.

Are there any questions? The starting time is _____. Begin working.

Day 5 Instructions

Good morning (afternoon).

Today only some of you will receive a test. Those of you who don't receive a test should work on the assignment given to you by your teacher.

(Distribute tests, pencils and rulers.)

Check your name. Write the name of the teacher and the class period where it says "teacher." Fill in your grade and the date.

(Write this information on the board as before.)

Raise your hand when you have finished the test. Begin working by reading the directions on the cover. Are there any questions?

(Collect all materials as students finish.)

Day 21 Instructions

Good morning (afternoon).

Today's testing will complete our study of geometry. I am back to test you in order to find out how much you remember from the lessons you completed three weeks ago.

(Distribute tests, pencils, and rulers.)

Does everyone have a test with his name on it? Do not open the test until I tell you to. Write the name of your teacher and the class period where it says "teacher." Fill in your grade and the date. (Put this information on the board.) Now look at the instructions.

(Read the instructions and do the examples.)

Are there any questions? Work carefully and do the best you can. When you are finished, raise your hand, then I will collect your materials. Begin working on the assignment your teacher has given you.

(Collect materials as students finish. After all materials have been collected, thank the students for their cooperation and pass out a piece of candy to each student. You may answer any questions about the experiment that the students may ask.)

GPO 827-442-3

Appendix C
Arithmetic Achievement Scores and Scores On Test BN for Each Subject

Student Number[a]	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score[b]
		CLASS 1		
1	6		16	
2	8	6.1	14	
3	2	7.7	17	20
6	2	4.6	14	12
7	6	5.9	15	
9	8	4.6	3	
10	2	7.1	19	19
11	7	10.6	17	17
12	5	7.7	26	22
13	5	7.1	26	27
14	5	6.3	28	28
15	5	6.6	21	18
17	1	7.6	28	
18	5	8.3	28	28
19	8	4.0	8	
20	4	7.4	28	
21	3	5.7	13	
22	2	8.0	15	20
23	1	12.2	27	
24	7	6.1	9	8
25	8	6.6	11	
26	4	6.6	27	
29	4		10	
		CLASS 2		
31	8	6.3	15	
34	8	7.1	18	
35	1	6.1	13	
36	8	4.6	10	
37	1		22	
38	6	11.5	28	
39	1	6.5	28	
41	4	4.4	25	

[a]Scores are given only for subjects who completed all lessons and tests.
[b]Retest scores are given only for subjects in the 1-day groups which were retested at 21 days.

Student Number	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score
		CLASS 2 (cont.)		
42	6	5.6	3	
44	3	6.8	16	
46	4	11.9	28	
47	7	6.8	22	8
48	4	5.7	9	
50	3	6.6	25	
51	2	6.1	15	14
53	3	8.3	14	
55	6	4.6	5	
56	2	3.6	9	9
57	2	8.0	28	28
58	4	4.6	16	
59	7	4.6	6	7
60	4	3.8	27	
62	4	4.9	23	
		CLASS 3		
65	5	4.0	9	4
66	3	4.4	8	
69	2	4.1	9	14
70	2	4.6	6	10
71	3	5.1	11	
72	3	4.9	9	
74	2	3.8	9	11
75	2	2.9	9	8
76	2	4.9	8	8
78	1	4.0	14	
79	3	3.6	10	
81	4	4.0	13	
82	2	5.6	8	13
83	7	4.2	10	6
84	3	3.4	7	
85	2	4.0	4	8
86	8	3.6	10	
87	1	5.4	16	
		CLASS 4		
90	5	5.9	14	7
93	3		21	
94	3	6.5	14	
95	1	6.6	18	
96	5	4.9	10	12
97	7	4.4	17	16
101	5	4.6	24	17
102	2	8.3	28	28
103	7	5.4	15	17
104	4	5.9	21	
105	2	4.2	13	13
107	8		15	
108	8	10.1	24	
110	7	6.5	14	15
111	4	7.7	28	

Student Number	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score
		CLASS 4 (cont.)		
112	3	6.8	20	
113	3	5.6	16	
114	6	4.0	9	
115	4	8.0	27	
116	7	4.9	16	17
117	5	9.6	28	25
118	5	4.6	28	18
119	5	6.6	24	14
120	2	7.1	26	27
121	1	4.0	12	
123	5	3.1	11	6
		CLASS 5		
125	4	9.1	23	
126	2	4.6	14	8
127	6		19	
128	2	12.2	28	28
129	1	5.9	20	
131	1	9.1	27	
132	4	4.2	22	
136	4	7.1	27	
137	3	4.0	13	
138	5	8.0	28	27
139	1	4.9	13	
140	4	8.3	27	
141	8	5.6	12	
142	2	5.4	13	11
143	4	6.6	24	
144	1	4.4	15	
145	8	8.0	16	
146	8	5.7	9	
148	8	6.6	8	
149	5	6.5	22	18
151	3	7.7	15	
152	1		28	
154	4	5.4	22	
155	4	5.1	13	
		CLASS 6		
156	2	6.1	9	12
157	3	5.7	9	
158	6	4.6	13	
160	6	6.8	20	
161	1	4.6	10	
162	7	5.6	10	11
164	2	6.8	24	27
166	7	8.0	9	8
167	8	6.5	14	
168	7	7.1	11	11
169	5	5.6	9	8
170	6	4.4	9	
172	6	6.6	9	

Student Number	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score
		CLASS 6 (cont.)		
173	6	3.6	12	
174	4	5.9	25	
175	1	10.1	21	
177	2	6.6	16	23
178	7	5.9	8	10
179	3	7.4	9	
180	6	7.1	25	
181	3	6.6	11	
182	1	5.6	12	
184	6	5.1	13	
186	8	6.8	3	
		CLASS 7		
187	7		11	10
188	6	4.2	13	
189	8	4.9	6	
191	5	5.4	12	14
192	6	6.5	15	
193	3		11	
194	6	8.0	14	
195	2	9.6	28	26
198	5	7.1	25	21
199	5	4.4	10	12
200	3	12.2	23	
201	7	6.3	12	19
202	6	3.1	16	
203	6	3.7	21	
207	2		18	15
208	7	5.7	7	6
209	5	4.2	11	10
210	8	4.0	7	
211	1		26	
212	1	6.8	26	
213	1	6.6	23	
214	6	4.9	23	
216	7	4.4	9	12
217	5	3.8	15	8
218	8		11	
219	2	8.3	19	13
335	2		23	21
		CLASS 8		
221	8	4.9	6	
222	7	6.6	8	4
223	3	5.1	11	
224	7		9	11
225	7	4.6	10	5
229	1	5.9	21	
230	4	4.6	13	
231	5	4.0	8	11
232	6	5.4	12	
233	8	4.2	8	

Student Number	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score
		CLASS 8 (cont.)		
234	6	4.0	5	
235	7	3.8	5	3
236	3	4.6	9	
237	2	4.2	9	10
238	4	3.6	4	
240	5	5.7	10	13
241	4	4.4	14	
242	7	4.9	7	8
		CLASS 9		
244	3	6.1	13	
245	1	8.3	27	
247	7	6.1	8	11
248	5	6.8	22	25
250	2		9	5
251	8	8.6	15	
252	6	6.1	20	
253	3	5.1	8	
254	8	4.6	10	
255	6	10.0	23	
257	7	4.6	5	13
259	7	5.6	7	9
260	2	4.9	10	9
262	3	8.6	23	
263	7	6.8	10	12
264	4	4.0	18	
270	7	7.4	22	22
272	4	6.6	26	
273	5	6.6	16	15
275	8	7.7	18	
		CLASS 10		
277	7	9.1	10	14
278	7	7.7	14	16
281	6	5.1	10	
282	2	6.4	16	16
283	5	7.4	14	8
284	1	6.6	23	
285	2	5.7	11	14
286	3	5.4	15	
288	1	6.1	10	
290	3	10.6	23	
291	4	6.1	27	
292	3	6.1	13	
293	8	11.6	16	
294	3	4.6	11	
295	6	8.6	17	
296	8	7.4	6	
297	4	6.8	28	
298	1	7.7	20	
299	4	9.1	27	
300	8	9.6	27	

Student Number	Treatment Number	Stanford Achievement Test Applied Arithmetic Score	Test Score	Retest Score
		CLASS 10 (cont.)		
301	6	6.5	13	
302	4	5.1	18	
305	3		10	
		CLASS 11		
307	8	5.7	9	
309	6	5.9	11	
311	6	6.3	13	
313	6	5.7	20	
314	6	6.6	24	
315	7	8.0	16	20
316	7	6.5	16	15
317	6	6.8	19	
318	8	12.5	21	
320	1	5.7	24	
321	5	5.7	13	11
322	2	6.5	17	13
323	1	4.9	9	
324	2	4.6	7	7
325	6	5.6	13	
326	3	5.7	18	
327	3	3.8	9	
329	4	6.1	26	
330	4	10.1	27	
331	1	5.6	17	
332	7	10.6	19	27
333	1	7.1	14	
334	2		7	6

GPO 827-442-2

National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
Universiy of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management
and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts